



# Australian Bureau of Statistics

## 2962.0 - Census Working Paper 93/3 - Posted-in Forms, 1991

Latest ISSUE Released at 11:30 AM (CANBERRA TIME) 06/01/1994

## Summary

### Main Features

### Census Working Paper 93/3

#### 1991 CENSUS DATA QUALITY: A STUDY OF POSTED-IN FORMS

Dietmar Kahles, Catriona Bate , Julie Evans

Population Census Evaluation August 1993

### CONTENTS

#### Introduction

#### Form return rates

Background

The sample

Mailback and Non-contact rates of return

Comparison with overseas experiences

#### Incidence of posted-in forms

1986 and 1991 Censuses

State/Territory counts

Smaller area counts

Early/late counts

#### Not stated rates

Comparison of posted-in and collected forms

Comparison of posted-in forms received early and posted-in forms received late

Comparison of forms resulting from Mailback and Non-contact situations

#### Response to Name (Question 2)

PES records

135 CD sample

The link to general data quality

#### Conclusion

#### Appendixes

Problems associated with the large number of posted-in forms received as identified by DPC and Census Field Organisation staff

Back of Census form

'Request to Return' Card

Posted-in form counts (1986/1991 Censuses) - Sources  
Occupied private dwellings, 1986 and 1991 Censuses, by State  
Comparison of not stated rates for posted-in and collected forms  
Comparison of not stated rates for posted-in forms received early and posted-in forms received late  
Comparison of not stated rates for Mailback and Non-contact posted-in forms

## **LIST OF TABLES**

- 1 Distribution of Census forms, 135 CD sample
- 2 Proportion of Mailback/Non-contact dummies raised, 135 CD sample
- 3 Rates of return for Mailback/Non-contact situations, 135 CD sample
- 4 Rates of return for Mailback/Non-contact situations, 135 CD sample vs Australia
- 5 Insertion rate of posted-in forms, 1986 and 1991 Censuses, Australia
- 6 Inserted posted-in forms by State, 1986 and 1991 Censuses
- 7 Inserted posted-in forms by State as a proportion of occupied private dwellings, 1986 and 1991 Censuses
- 8 Inserted posted-in forms by State, Capital city Statistical Divisions and 'Other' Statistical Divisions, as a proportion of occupied private dwellings, 1991 Census
- 9 Proportion of forms received late
- 10 Variables where not stated rates were higher for posted-in forms
- 11 Variables where not stated rates were lower for posted-in forms
- 12 Response to 'Name' for posted-in and collected forms according to the 1991 PES
- 13 Response to 'Name' for Mailback and Non-contact forms, 135 CD sample
- 14 1991 Census not stated rates by response to 'Name' according to the 1991 PES

## **INTRODUCTION**

In the first few weeks of processing staff at the DPC were confronted with the arrival of an unprecedented number of Census forms by mail. The sheer volume of forms arriving every day greatly exceeded all expectations and posed a variety of logistical problems for management. The main concern expressed was about the impact on overall data quality as mailed forms were felt to generally include poorer quality data. This experience also showed the need to analyse the issue of posted-in forms in detail to enable the development of strategies for future Censuses.

One can only speculate about the exact reasons for the large increase in posted-in forms for the 1991 Census. It seems that privacy concerns were of greater importance to people in the 1991 Census than they were in 1986. This observation is backed up by the contents of notes attached to Census forms received at the DPC and by comments written on Census forms. Most of these comments expressed concerns about invasion of privacy, confidentiality or 'big brother' government. To what extent the public was influenced by some ill founded negative publicity in the media before Census night cannot be ascertained. Some people simply used the mailback facility to voice general comments about Government, Public Service, 'big business', religious beliefs etc.

Some problems associated with the large number of posted-in forms received which were identified by DPC and Census Field Organisation staff are discussed in Appendix 1. This report does not attempt to give solutions for these problems, but focuses on the data quality implications. Both return rates (how many forms were actually returned for every dummy form raised in the field) and variable not stated rates are used as indicators of data quality.

Before addressing these aspects in detail, it is necessary to give a definition of posted-in forms. Posted-in forms are defined as any Census form (household or personal) returned by mail rather than collected by a Census collector. A Census form was identified as posted-in

by DPC staff OMR-marking one of the three categories (posted-in early - complete, posted-in early - blank, posted-in late) of the 'MF' (mailed form) box on the back of the Census form. The distinction between 'early' and 'late' was made for the purposes of the Post Enumeration Survey (PES) and was determined by the commencement date of the PES (about three weeks after Census night). Only forms identified this way were able to be recognised as posted-in when interrogating the IFURF (Interim Final Unit Record File). Indications are that only very few posted-in forms were not marked correctly. Also, approximately 2% of posted-in forms received (about 4,000 forms) were deliberately not inserted in CD packs (and were thus not recorded on the IFURF) for reasons explained in **Incidence of posted-in forms**.

## **FORM RETURN RATES**

### **Background**

One of the aims of this study was to determine how many people who told their collector they would mail their form (mailback) actually did so, in contrast to those who could not be contacted by collectors. Rates of return give an indication of how much data might be missing as a result of mailback procedures.

There are two situations which may lead to a Census form being posted-in:

- A collector contacts the householder to collect the form. Mailback procedures should only be followed if it will convert/avoid a potential Refusal situation and only after the householder has been offered a Gold Privacy envelope. The collector issues an envelope with a mailing address sticker attached and prepares a dummy form with the 'DF' (Dummy Form) box on the back of the form marked 'MB' (Mailback). (A copy of the back of the 1991 Census form can be found in Appendix 2.) In the Collector Record Book he/she will write 'MB' in the remarks column. This situation is called 'mailback' or 'MB'.
- A collector fails to contact the householder to collect the form. If up to five follow-up calls remain unsuccessful, the collector will leave another household form with a 'Request to Return' card (see Appendix 3 for a copy of this card) and a privacy envelope with a mailing address sticker attached. The collector will then prepare a dummy form and mark the appropriate box with the 'NC' (Non-contact) code. In the Collector Record Book he/she will write 'NC' in the remarks column. Some householders however did return the form in private envelopes, posting it to the nearest ABS office or to the DPC. This situation is called 'non-contact' or 'NC'.

### **The sample**

Initially, 139 CDs from around Australia were randomly selected by Statistical Support for this study. Four of the CDs selected (three RAIF area CDs, and one CD for which a Collector Record Book could not be found) were excluded from analysis, leaving 135 CDs for the study. Within these CDs, a further selection of records was made in the interests of efficiency and to ensure that sufficient numbers of posted-in forms were included. Every posted-in form from these 135 CDs was included to ensure the number was large enough for analysis. 10% of collected forms were selected at random for inclusion in this study as a control group (as in the Contact at Delivery study).

This selection procedure will need to be kept in mind when in the remainder of the report reference is made to the '135 CD sample'. The result was a selection of 4,353 forms. 17% of these (724 forms) were posted-in forms, higher than the incidence of posted-in forms in the

general population (about 3%) due to the way forms were selected. Table 1 shows the distribution of records resulting from this selection. Posted-in forms are the focus of this analysis.

**Table 1: Distribution of Census forms, 135 CD sample**

Category	Households	
	No.	%
Collected forms	3,629	83.4
Posted-in forms	724	16.6
Total	4,353	100.0

To determine the number of mailback and non-contact dummies raised in the field the Remarks column of each Collector Record Book for all 135 CDs was examined and the incidence of mailback and non-contact was tallied. A total of 1,047 mailback and non-contact dummies were identified this way. About three quarters of these dummies were recorded as mailback, and this means the collector was able to contact the householder in the majority of cases.

**Table 2: Proportion of Mailback/Non-contact dummies raised, 135 CD sample**

Category	Dummies raised	
	No.	%
Mailback	806	77.0
Non-contact	241	23.0
<b>Total</b>	<b>1,047</b>	<b>100.0</b>

### Mailback and Non-contact return rates

**Table 3: Rates of return for Mailback/Non-contact situations, 135 CD sample**

Category	Dummies raised		Forms received	Rate of return
	No.			
Mailback	806		No. 644	% 79.9
Non-contact	241		80	33.2
<b>Total</b>	<b>1,047</b>		<b>724</b>	<b>69.2</b>

Table 3 shows the number of dummy forms replaced by posted-in forms received at the DPC. Forms were far more likely to be posted when the collector was able to contact the householder on collection (four returns for every five dummies) than in cases of non-contact (one return for every three dummies). Overall, about 70% of forms that were not collected (but could not be classed as Refusals) were posted by householders. If the fact that not all

posted-in forms were used (they were not inserted if they were blank or arrived too late or contained obviously untrue or unidentifiable information) is taken into account, the proportion of forms received by mail would be slightly higher.

An overall return rate which combines mailback and non-contact dummies can also be calculated for Australia. A comparison of these rates for the 135 CD sample and Australia reveal that the estimate for the sample is very close to the rate for Australia. However, the fact that the rate for Australia is even lower than the already low return rate for the sample is disappointing.

**Table 4: Rates of return for Mailback/Non-contact situations, 135 CD sample/Australia**

Source	Dummies raised		Forms received		Rate of return
	No.		No.		
135 CD sample	1,047		724		69.2
Australia	262,670		171,199*		65.2

\* For comparison purposes, 4,000 posted-in forms received but not inserted have not been included.

### Comparison with overseas experiences

Interestingly, the return rate achieved for posted-in forms in Australia (about 70%) is approximately in line with return rates for overseas (mailback) Censuses. This overall return rate is likely to be similar to the situation after first follow-up in these Censuses, given that households for which mailback or non-contact dummies are raised are reminded in person (mailback) or by card (non-contact) to complete and return Census forms. Although direct comparison is difficult, it is interesting to note that in Canada in 1986 intensive follow-up was needed to achieve an 87% return rate. The return rate before follow-up (but after an initial request) in the US in 1990 was only 65%.

### INCIDENCE OF POSTED-IN FORMS

#### 1986 and 1991 Censuses

Table 5 shows the total number of posted-in forms received for the 1986 and 1991 Censuses, and the number of forms inserted in CD packs at the DPC. Forms returned by post were not inserted if they were blank or contained no useful information.

**Table 5: Insertion rate of posted-in forms, 1986 and 1991 Censuses, Australia**

Category	1986 Census		1991 Census	
	No.	%	No.	%
Inserted	51,320	89.6	171,199	97.7
Not inserted	3,251	5.7	4,000*	2.3
Unknown	2,690	4.7	-	-
<b>Total</b>	<b>57,261</b>	<b>100.00</b>	<b>175,199</b>	<b>100.00</b>

\* estimated

To derive the figures in Table 5 it was necessary to access different sources for the 1986 and 1991 Censuses. 1986 data originated from tally forms used by Special Projects at the DPC, whereas 1991 data originated from running a SAS program against the posted-in form field on the IFURF. Additional evidence was provided by ex-DPC staff. The problems experienced in deriving the data are discussed in detail in Appendix 4.

In 1986, about 57,000 posted-in forms were received by the DPC. Of these 57,000 forms about 3,250 are known to have been blank and were thus not inserted in CD packs to be processed. This represents a proportion of about 5.7% not inserted in CD packs. The 'Unknown' category represents forms originally thought to be refusals which were later found to be posted-in forms. It is not known how many of these forms were blank and were therefore not inserted. The proportion of forms not inserted is therefore likely to have been even greater than 5.7%.

In 1991, about 175,000 posted-in forms were received by the DPC. About 4,000 forms are estimated to not have been inserted in CD packs. The large majority of these 4,000 forms were blank. A much smaller number were not inserted because they either contained obviously untrue and/or incomprehensible information, or had arrived too late at the DPC to be inserted at any stage of processing. This means about 2.3% of posted-in forms were not inserted in CD packs. Table 5 shows that the number of posted-in forms has tripled between the 1986 and 1991 Censuses. Assuming most of the forms not inserted were blank, it can also be seen that the proportion of blank forms decreased between 1986 and 1991.

### State/Territory counts

Table 6 shows the number of inserted posted-in forms by State for the 1986 and 1991 Censuses, and the proportion of inserted posted-in forms. New South Wales, Queensland and South Australia increased their share of posted-in forms, while all other States and Territories showed decreases. The Australian Capital Territory, Tasmania and the Northern Territory all decreased by between 30% and 70%. The lower intercensal increase in the number of posted-in forms for the smaller States and Territories may have been partly a result of the effect of the absence of 4,000 posted-in forms which were not inserted in CD packs, as anecdotal evidence suggests that they mainly originated from the smaller States and Territories.

**Table 6: Inserted posted-in forms by State, 1986 and 1991 Censuses**

State	1986 Census		1991 Census		Ratio of 1991 to 1986 number
	No.	%	%		
	of posted-in forms		of posted-in forms		
New South Wales	17,792	34.7	61,965	36.2	3.5
Victoria	16,107	31.4	52,247	30.5	3.2
Queensland	6,542	12.8	25,518	14.9	3.9
South Australia	3,975	7.8	14,584	8.5	3.7
Western Australia	4,274	8.3	12,576	7.4	2.9

Tasmania	1,101	2.2	1,860	1.1	1.7
Northern Territory	799	1.6	775	0.5	1.0
Australian Capital Territory	730	1.4	1,674	1.0	2.3
<b>Australia</b>	<b>51,320</b>	<b>100.0</b>	<b>171,199</b>	<b>100.0</b>	<b>3.3</b>

Disregarding the potential effect of the 4,000 posted-in forms which were not inserted there seems to be a difference in the size of the intercensal increase in the number of posted-in forms between the five biggest States and the three smallest States or Territories. While the number of mailed forms in New South Wales, Victoria, Queensland, South Australia and Western Australia tripled or more than tripled, it increased to a lesser extent in the Australian Capital Territory and Tasmania, and remained stable in the Northern Territory. This cannot be explained by differences in intercensal population growth rates between the States and Territories, as the Australian Capital Territory and the Northern Territory experienced amongst the biggest growth rates, but showed a comparatively smaller increase or even a small decrease in the number of posted-in forms.

Table 7 shows the proportion of posted-in forms by State as a percentage of occupied private dwellings (see Appendix 5 for data on the number of occupied private dwellings by State). This indicates differences in the distribution of posted-in forms between the States and Territories. Again, a difference between the bigger and smaller States and Territories is noticeable. In Tasmania, only about one in hundred households (1.1%) posted their form, whereas in Victoria about one in thirty households (3.5%) decided to do so. Intercensal increases in the proportion of posted-in forms range from 0.4% (Tasmania) to 2.3% (Victoria). The Northern Territory was the only State or Territory not showing an increase in the proportion of posted-in forms. Because the number of posted-in personal forms is not known, it is not possible to say exactly how many occupied private dwellings were represented by posted-in forms. However, mailback of personal forms can only arise in very unusual circumstances and the number is thought sufficiently small in both Censuses for proportional changes as shown in Table 7 not to be affected.

**Table 7: Inserted posted-in forms by State as a proportion of occupied private dwellings, 1986 and 1991 Censuses**

State	1986 Census	1991 Census	Percentage Points Difference
	%	%	
New South Wales	1.0	3.1	2.1
Victoria	1.2	3.5	2.3
Queensland	0.8	2.5	1.7
South Australia	0.8	2.8	2.0
Western Australia	0.9	2.3	1.4

Tasmania	0.7	1.1	0.4
Northern Territory	1.9	1.5	-0.4
Australian Capital Territory	0.9	1.8	0.9
<b>Australia</b>	<b>1.0</b>	<b>2.9</b>	<b>1.9</b>

### Smaller area counts

Within States and Territories there tended to be differences in the incidence of posted-in forms between Statistical Divisions and Subdivisions in the 1991 Census. Capital City Statistical Divisions showed higher proportions of posted-in forms than Statistical Divisions representing the remainder of a State or Territory.

The incidence of posted-in forms was calculated as the number of posted-in forms over the number of occupied private dwellings for each Statistical Subdivision. The resulting data were then aggregated to Statistical Divisions and further to Statistical Divisions representing Capital Cities and 'Other' State areas. These two groups were identified by their ASGC codes (05 for Capital City Statistical Divisions).

This analysis concentrated on the comparison of Capital City and 'Other' SDs as estimates for smaller areas were thought to be subject to an unacceptably high level of sampling error due to the small sample sizes of posted-in forms involved.

Table 8 compares the incidence of posted-in forms in Capital City Statistical Divisions with the incidence of posted-in forms in 'Other' Statistical Divisions.

**Table 8: Inserted posted-in forms by State, Capital City Statistical Divisions and 'Other' Statistical Divisions, as a proportion of occupied private dwellings, 1991 Census**

State	Capital City Statistical Division %	'Other' Statistical Division %	Percentage Points Difference
New South Wales	3.6	2.2	1.4
Victoria	4.2	1.9	2.3
Queensland	2.9	2.2	0.7
South Australia	2.9	2.6	0.3
Western Australia	2.7	1.2	1.5
Tasmania	1.2	1.1	0.1

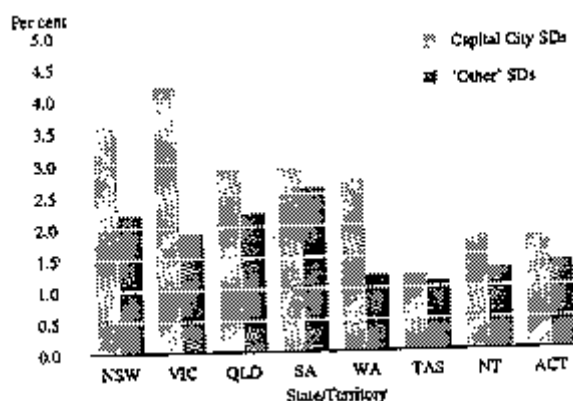


Northern Territory	1.8	1.3	0.5
Australian Capital Territory*	1.8	1.4	0.4
<b>Australia</b>	<b>3.4</b>	<b>2.0</b>	<b>1.4</b>

\* The Canberra Statistical Division incorporates all Canberra town centres. 'Other' Statistical Divisions include the rest of the ACT and Jervis Bay.

For all States and Territories the incidence of posted-in forms seemed to be higher in Capital Cities than in the rest of the State or Territory. The figure below illustrates those differences in detail. The more populous States showed generally larger differences in the incidence of posted-in forms. Differences tended to be greatest in Victoria (2.3%), Western Australia (1.5%) and New South Wales (1.4%), and smallest in Tasmania (0.1%) and South Australia (0.3%).

### INCIDENCE OF POSTED-IN FORMS Capital City versus 'Other' Statistical Divisions



Note: The incidence of posted-in forms is shown as a proportion of occupied private dwellings

In Victoria the incidence of posted-in forms was more than two times greater for the Melbourne Statistical Division than for the State's other Statistical Divisions. Melbourne also showed the highest incidence of posted-in forms (4.2%) of any Capital City Statistical Division in Australia.

The proportion of posted-in forms within Statistical Subdivisions (SSDs) varied little in Capital City SDs but varied greatly in 'Other' SDs. The level of posted-in forms for Capital City SSDs was consistently high but there was little variation (Perth showed the least variation, from 2.3% to 3.2%). In contrast, there was larger variation in 'Other' SDs, with the greatest variation for SSDs recorded in South Australia (from 0.7% to 5.3%). It should be remembered, however, that these results are subject to substantial sampling error due to the small sample sizes involved.

### Early/Late counts

The number of forms received early (in the first three weeks after Census night) versus the number of forms received late is shown in Table 9 below. About one fifth of all posted-in forms, for both the 1986 and 1991 Censuses, arrived late.

This could be an indication of the relatively minor importance of deliberately delaying to return the Census form as a form of protest. It appears more likely that people who wanted to register their protest against the Census (for whatever reason) either did not return their form at all (disguised refusals) or decided not to answer certain questions. Supporting evidence for this assumption together with the effect of mailing the Census form late on data quality will be discussed in **Not Stated Rates**.

Other problems associated with late forms were that a proportion of these forms could not be inserted in CD packs (especially forms from States processed early in the cycle) and that they needed to be excluded from the Post Enumeration Survey which commenced about three weeks after Census night.

**Table 9: Proportion of forms received late**

Category	1986 Census		1991 Census	
	No.	%	No.	%
Early	46,051	80.4	136,240	79.6
Late	11,210	19.6	34,959	20.4
<b>Total</b>	<b>57,261</b>	<b>100.0</b>	<b>171,199</b>	<b>100.0</b>

Note: 1991 data excludes forms which were not inserted.

## NOT STATED RATES

Not stated rates for specific variables or questions are a basic indicator of data quality. This analysis will make three comparisons between not stated rates by variable using different sets of data in order to draw some conclusions about data quality for posted-in forms. First, not stated rates for variables for records from posted-in forms will be compared to those from collected forms. Posted-in forms will then be analysed in more detail to determine whether there was a difference in not stated rates between:

- posted-in forms received early and posted-in forms received late; and
- the two posted-in form situations 'mailback' and 'non-contact on collection'

Although the not stated rates quoted here do resemble non-response rates (as they appear in Fact Sheet 11.0) they should not be treated as equivalent. All not stated rates in this analysis originate from IFURF data and there are some conceptual differences as well as the potential for differences due to additional editing and reformatting taken account of in FURF data (the source of Fact Sheets). Changes in the definition of the applicable population affect a number of variables, particularly Year of Arrival, Proficiency in English, Type of educational institution and the Qualification variables. However, these differences are not expected to affect the comparisons made.

There are also some differences in the list of variables used in this analysis. Most of the variables taken off the IFURF relate directly to questions as they appear on the Census form. A few variables (for example, Work destination zone or Occupation) were derived from more than one question. Other derived variables, such as Labour force status, were not available from the IFURF used. However, the labour force questions Full/Part-time job,

Looked for work and Job last week were available and are included in this analysis. Some other variables were not available or were excluded due to doubts about the quality of the data at that stage. These include Qualification (highest field), Rent (weekly) and Marital Status (imputation rate).

While IFURF data for Australia were the source for most of the analysis, IFURF data corresponding to records used in the 135 CD sample (see 2.2 for sample details) were separately analysed in order to compare the two posted-in form situations 'mailback' and 'non-contact'. In addition, it was necessary to clerically examine forms used in the 135 CD sample before interrogating the IFURF to derive not stated rates from the corresponding records.

Records from dummy forms were excluded from analysis. Apart from inflating not stated rates because most of the data are missing, the main problem with dummy forms is that they represent forms which were not received at the DPC, and which therefore could not be classified as either collected or posted-in forms.

The absence of the 4,000 or so posted-in forms which were not inserted during any stage of processing will have the effect of slightly underestimating not stated rates for posted-in forms in part 1 of the analysis, and slightly reducing the size of the sample of posted-in forms used in part 2 of the analysis. Part 3 is unlikely to be affected very much because it relies on a random sample of only 135 CDs from all over Australia.

### **Comparison of posted-in and collected forms**

Not stated rates tended to be higher for posted-in forms for most variables. See Appendix 6 for not stated rates by variable separately for posted-in and collected forms. The overall pattern of not stated rates is very similar for both sets of data and also quite similar to that found in FURF data.

The differences in not stated rates range from 0.1 percentage points for Birthplace to 6.7 percentage points for Religion. Tests of statistical significance (comparison of means at a 95% confidence interval) indicated that the difference between not stated rates for posted-in and collected forms was significant for all variables except Occupation. The large number of records made the test more powerful, enabling the detection of small differences as significant.

The overall percentage point differences in not stated rates between records for collected and posted-in forms are not great. In this situation, the absence of records from posted-in forms which were received but not inserted in CD packs may be important because this means that not stated rates calculated for posted-in forms are slightly understated. Although the absent posted-in forms represent only about 2% of all posted-in forms (as shown in Table 5), if these forms were able to be included in the data not stated rates for posted-in forms would have been higher because of the high proportion of the missing forms that were blank or nearly so.

Not stated rates for posted-in forms would increase by up to two percentage points if the absent posted-in forms were able to be included. The impact would be greatest for variables which are applicable to the whole population, and the percentage points differences between not stated rates for collected and posted-in forms would be affected. In most cases the extent of the differences would be increased.

The variables in Table 10 below are cases where not-stated rates for posted-in forms were higher than not stated rates for collected forms. They represent about two thirds of the variables studied. Only five show a percentage points difference of more than 1%. While the

differences shown do reach statistical significance, in all cases this difference would be significantly increased were the absent forms to be included due to the relatively small size of the differences shown below.

**Table 10: Variables where not stated rates were higher for posted-in forms**

Variable	Not stated rates		Percentage points
	Collected forms %	Posted-in forms %	Difference
<b>Person variables</b>			
Age	0.7	0.9	0.2
State of usual residence (Census night)	4.2	5.0	0.8
State of usual residence (1 year ago)	3.6	3.9	0.3
Usual address 5 years ago	2.5	3.0	0.5
Birthplace	0.9	1.0	0.1
Aboriginal/TSI	3.2	3.4	0.2
Birthplace of father	1.3	1.6	0.3
Birthplace of mother	2.0	2.3	0.3
Australian citizenship	2.0	2.2	0.2
Religion	8.8	15.5	6.7
Language	1.5	1.7	0.2
Proficiency in English	7.6	8.7	1.1
Individual income	7.9	11.1	3.2
Looked for work	11.8	12.5	0.7
Industry sector	7.8	8.3	0.5
Industry	9.0	9.5	0.5
Destination zone	14.2	15.0	0.8
<b>Dwelling variables</b>			
Number of motor vehicles	1.5	1.7	0.2
Number of bedrooms	1.5	2.0	0.5

Nature of occupancy	1.4	1.9	0.5
Landlord	17.0	17.8	0.8
Furnished/Unfurnished	20.8	21.6	0.8
Housing loan repayments	3.3	4.4	1.1
Structure of dwelling	0.5	5.4	4.9

For other variables, the impact of the additional 4,000 posted-in forms would perhaps change the direction of the percentage points difference in not stated rates. Variables appearing in Table 11 below have lower not stated rates for posted-in forms than collected forms. The addition of the absent forms would probably result in slightly higher or nearly equal not stated rates for posted-in forms compared to collected forms for these variables. The only variables for which collected forms might continue to show a higher not stated rate are Year of qualification and possibly Job last week.

**Table 11: Variables where not stated rates were lower for posted-in forms**

Variable	Not stated rates		Percentage points
	Collected forms %	Posted-in forms %	Difference
<b>Person variables</b>			
Year of arrival	5.5	5.2	0.3
Full/Part-time student	2.5	2.2	0.3
Type of educational institution	13.3	13.0	0.3
Age left school	6.7	6.2	0.5
Qualification, (highest) level	15.2	14.4	0.8
Qualification, year obtained	17.4	15.4	2.0
Full/Part-time job	4.6	4.3	0.3
Job last week	11.1	9.6	1.5
Hours worked	7.2	7.0	0.2
Method of travel to work	6.5	5.8	0.7

The most remarkable difference in not stated rates was for the Structure of dwelling variable where the not stated rate (5.4%) was more than nine times higher for posted-in forms than for collected forms. In comparison, the not stated rate for the Structure of dwelling variable derived from final data is 0.8% for Australia (Fact Sheet 11.0). This appears to be a strong indication of collector practice departing from correct procedure. It seems that in many cases collectors may not have completed this question before delivery as instructed in the Collectors Manual, but instead have completed it only after collection of the form.

Higher not stated rates for posted-in forms could, for some variables, be interpreted as a registration of protest against a perceived invasion of privacy. Some support for this argument may be seen in the fact that the two largest differences in not stated rates, apart from Structure of dwelling, were for the traditionally contentious variables Religion and Income. It seems that a significant number of people who chose to post their form registered their concern about privacy issues by not responding to sensitive questions.

### **Comparison of posted-in forms received early and posted-in forms received late**

Appendix 7 shows the not stated rates for posted-in forms received early and posted-in forms received late. Comparing these two categories of posted-in forms should reveal any link that might exist between time taken to return a Census form by post and not stated rates (data quality). Forms received in the first three weeks after Census night are referred to as received early, forms received thereafter are referred to as received late. This reflects the requirements of the PES which commenced three weeks after Census night and included only posted-in forms received before its commencement. As noted earlier, about 80% of all posted-in forms were received early.

The data was again derived from the IFURF (Australia) by tallying the number of not stated codes for the posted-in form categories 'received early' and 'received late', the number of applicable records for each variable, and calculating the not stated rate as a proportion of not stated codes over total applicable records. Although the 'received late' category might be expected to be affected more than the 'received early' category by the absence of the 4,000 posted-in forms, the fact that forms were inserted relatively late in processing and were not selected for insertion on this basis means that this comparison is unlikely to be affected by this factor.

Tests of statistical significance (comparison of means at a 95% confidence interval) indicated that the difference between not-stated rates for posted-in forms received early and posted-in forms received late was significant for all but three variables (Income, Landlord, Furnished/Unfurnished). Due to the large number of records small differences were able to be detected as significant.

The pattern of not stated rates is very similar for posted-in forms received early and posted-in forms received late. For those variables where there was a significant difference all except Religion showed higher not stated rates for posted-in forms received late. Differences range from -1.0 percentage points (Religion) to 3.4 percentage points (Type of educational institution).

That not stated rates for almost all variables are significantly higher for posted-in forms received late may indicate that the commitment to answer Census questions decreases slightly over time, although the overall response rates are still relatively high. There is no particular group of variables showing higher not stated rates than other groups.

Two interesting results for posted-in forms received late are apparent for Religion and Income (which showed the highest not stated rates after Structure of dwelling in the

posted/collected forms comparison). Religion was the only variable with a lower not stated rate and Income was the only variable for which there was no significant difference from the rate for posted-in forms received early. That these sensitive variables did not show higher not stated rates might suggest that privacy may not be such an issue for the group of people posting their form late.

### **Comparison of forms resulting from Mailback and Non-contact situations**

In order to determine possible differences in not stated rates for mailback and non-contact, posted-in forms from the sample of 135 CDs were analysed. They were classified as mailback or non-contact by examining Collector Record Books, and the corresponding records on the IFURF were interrogated to obtain not stated rates for variables according to posted-in form type.

Appendix 8 contains not stated rates for mailback and non-contact forms. The very low incidence of non-contact posted-in forms presented a major problem for this analysis. Only variables with a tally of more than three not stated codes per variable (mailback or non-contact forms) were included to present more meaningful results. However, the numbers involved are still very small (especially for non-contact forms), and results should only be regarded as indicative of possible differences.

It seems that for the majority of variables not stated rates for posted-in forms were higher in situations where there was contact with a collector at collection (mailback). This is a somewhat surprising result and could possibly add weight to the opinion voiced by some that a proportion of posted-in forms should be regarded as disguised refusals. The result is also reminiscent of the conclusion in the Contact at Delivery study that higher not stated rates were found where there was contact at delivery than where there was none. In both cases there needs to be an examination of the characteristics of the household, and characteristics of persons contacted to evaluate this phenomenon.

### **RESPONSE TO NAME (QUESTION 2)**

Response to 'Name' can be regarded as an important supplementary indicator of overall data quality. Anecdotal evidence has previously been claimed to show that persons who include their name on the Census form also tend to complete the remainder of the form more thoroughly. Not giving a name may also highlight possible concerns about privacy and confidentiality issues. The following questions were targeted in this study:

- How does the proportion of records on posted-in forms with no response to 'Name' compare with the proportion of records on collected forms without a name?
- Within the posted-in forms group, was there a difference in response to 'Name' between mailback and non-contact (at collection) forms?
- Did Census records without a name tend to have higher question not stated rates than records with a name?

The level of response to Question 2 (Name) on the Census form could only be directly measured by clerical analysis since it is not captured during Census data processing and is therefore not recorded on Unit Record Files. However, it is possible to get an indication of the quality of response to 'Name' from the sample of forms used in the PES (Post Enumeration Survey) as the presence or absence of 'Name' is recorded on the PES file. The main source for this investigation was the 1991 Census PES file. To separately identify non-response for posted-in form mailback and non-contact cases the 135 CD sample was analysed.

### **PES records**

The following table was constructed using all persons on the PES file who were matched to the Census form corresponding to the PES address. A separate match code identified cases where 'Name' was missing on either the PES or the Census record. If the PES record was complete for that person then this match code would indicate that 'Name' was missing from the Census record. If the PES record was not complete, it could not be determined if 'Name' was omitted from the PES or the Census record. The latter cases are shown in the table as 'Unresolved'.

**Table 12: Response to 'Name' for posted-in and collected Census forms according to the 1991 PES**

Category	Posted-in forms		Collected forms		Total forms	
	No. of records	%	No. of records	%	No. of records	
Name given	1,730	94.4	85,535	99.1	87,265	99.0
Name missing	59	3.2	582	0.7	641	0.7
Unresolved	43	2.4	199	0.2	242	0.3
<b>Total</b>	<b>1,832</b>	<b>100.0</b>	<b>86,316</b>	<b>100.0</b>	<b>88,148</b>	<b>100.0</b>

Indications are that the proportion of people mailing their Census form who did not give their name was about four times higher (3.2%) than that of people whose forms were collected (0.7%). This difference was found to be statistically significant. Evidence from staff involved in the conduct of the PES suggests there were very few PES records with 'Name' missing, thus most of the 'Unresolved' cases in the above table would involve Census records with 'Name' missing. The proportion of records with no response to 'Name' could therefore be as great as 5.6% for posted-in forms and 0.9% for collected forms (ie. about six times higher for the former).

While non-response to Question 2 ('Name') is apparently higher for posted-in forms, 95% or more of records on these forms still contain names. This may indicate that privacy concerns of people mailing their Census form were not as important an issue as previously thought.

### 135 CD Sample

Table 13, derived from a clerical examination of posted-in forms in the 135 CD sample (coincidentally the total number of records on posted-in forms corresponds closely to that shown in Table 12), indicates possible differences in non-response between the two posted-in form situations, mailback and non-contact. However, the low incidence of non-contact forms and of records where names were missing means that tests of significance are invalid, and the results shown below should therefore be treated with caution.

**Table 13: Response to 'Name' for mailback/non-contact forms, 135 CD sample**

Category	Mailback		Non-contact		Total posted-in forms	
	No. of records	%	No. of records	%	No. of records	%



Name present	1,644	95.8	114	98.3	1,758	95.9
Name missing*	73	4.2	2	1.7	75	4.1
<b>Total</b>	<b>1,717</b>	<b>100.0</b>	<b>116</b>	<b>100.0</b>	<b>1,833</b>	<b>100.0</b>

\* Includes partial response (ie Surname or Christian name only)

Overall non-response to 'Name' (4.1%) again appears to be quite low. The data are fairly consistent with the results from the PES file analysis. The figures suggest slightly higher non-response to the 'Name' question for people requesting a mailback envelope compared to people where there was no contact at collection. Perhaps people who actively requested to make use of the mailback facility had a greater reluctance to include their name on the Census form (possibly another way of expressing concern about a perceived invasion of privacy). People who posted their form because no contact was made at collection seemed to have a similar response rate to 'Name' as people whose forms were collected (98.3% of the former as compared to 99.1% of the latter - see also Table 12).

### The link to general data quality

Table 14 compares the data quality of Census records without names to that of records with names. Anecdotal evidence from many sources suggests that where a respondent writes their name, responses to subsequent questions on the Census form tend to be more accurate. It was therefore suspected that the absence of names would be linked to generally higher non-response rates for other questions.

PES data were used for analysis. The same assumptions apply here as in Table 12 where PES data were also used. The variables analysed were Age, Birthplace, Marital Status and Aboriginal Origin, as these are the only variables for which Census responses were recorded on the PES file.

**Table 14: 1991 Census not stated rates by response to 'Name' according to the 1991 PES**

Variable Name given	Name missing	Unresolved	Total	
<b>Age</b>				
Stated	86,826	626	240	87,692
Not stated	439	15	2	456
Not stated rate (%)	0.5	2.3	0.8	0.5
<b>Birthplace</b>				
Stated	85,992	610	232	86,834
Not stated	1,273	31	10	1,314

Not stated rate (%)	1.5	4.8	4.1	1.5
<b>Marital status</b>				
Stated	82,532	585	220	83,337
Not stated	4,733	56	22	4,811
Not stated rate (%)	5.4	8.7	9.1	5.5
<b>Aboriginal origin</b>				
Stated	84,213	593	228	85,034
Not stated	3,052	48	14	3,114
Not stated rate (%)	3.5	7.5	5.8	3.5

Although absolute numbers for the categories 'Name missing' and 'Unresolved' are small compared to numbers for the category 'Name given' it can be concluded that for all variables, the not stated rate was higher for persons who did not include their name on the Census form than for those who did. As mentioned before, it is likely that most of the persons classed as 'Unresolved' omitted their name from the Census form. If this category is combined with those for whom it is known that name was missing, then the not stated rates are 1.9% (Age), 4.6% (Birthplace), 8.8% (Marital status), and 7.0% (Aboriginal origin). This means the not stated rates still remain higher for those who did not include their name on the form.

The results indicate that higher question response rates are linked to the inclusion of a person's name on the Census form. Given that people who post their form have a slightly higher tendency to omit their name from the Census form, this may have affected, albeit to a very small extent, overall Census data quality.

In the 1991 Census, only about 3% of forms were posted-in and 95% or more of these records contained names. Although the above analysis indicates the quality of response in records without names is likely to be worse than for other Census records, the number of such records is too small to have much effect on overall data quality.

## CONCLUSION

About three times as many posted-in forms were received for the 1991 Census as for the 1986 Census. They represented about 3% of all occupied private dwellings. There was a higher incidence in the more populous States and, within States and Territories, in Capital city Statistical Divisions.

The two main problems associated with posted-in forms in the 1991 Census were a **low overall return rate** and **higher not stated rates** for the majority of variables for records from posted-in forms.

Only about 70% of all Census forms that should have been returned by mail were actually

received at the DPC. This includes households requesting to mail the Census form at collection (Mailback) as well as households that were not able to be contacted at collection (Non-contact). More than three quarters of posted-in form situations were Mailback cases. The rate of return was higher for households contacted at collection (80%) than for households where there was no contact at collection (33%). Overall, this low rate of return resulted in the complete loss of almost one third of potential data from posted-in forms.

For the majority of variables not stated rates were higher on posted-in forms than on collected forms. Differences in not stated rates between posted-in and collected forms were generally not large, but all except one variable reached statistical significance. The most noteworthy results were:

- The not stated rate for the Structure of Dwelling variable (5.4%) was more than nine times higher for posted-in forms than for collected forms. As this variable is a collector-marked question the large difference in not stated rates suggests that collectors tended not to complete this question before delivery as instructed.
- The largest differences in not stated rates between posted-in and collected forms were for the two variables Religion (15.5% on posted-in forms and 8.8% on collected forms) and Income (11.1% and 7.9% respectively). It might be surmised that there is a link between people's reluctance to answer these questions and the use of postal return.

It seems that in cases where forms were not collected (excluding outright refusals) a distinction can be made between households which were generally willing to participate in the Census and others which were not.

Households returning Census forms by mail were largely co-operative. Although these households appeared to be generally willing to participate in the Census, they may have been more concerned about a perceived invasion of privacy or possible breaches of confidentiality than people whose forms were collected. The traditionally sensitive variables Religion and Income showed significantly higher not stated rates on posted-in forms than collected forms, whereas other variables showed only slightly higher not stated rates on posted-in forms. Interestingly, 95% of persons in households posting their Census form still put a name on the form. This might be an indication that they were quite willing to identify themselves knowing the collector would not see their name.

Only about 2% of all posted-in forms received at the DPC were blank or contained useless information. It appears that the majority of people who were not willing to participate in the Census chose to express that by not returning the Census form at all rather than by returning it blank or containing useless information. As mentioned before, only about seven out of ten Census forms that should have been returned by mail were actually returned. This was despite either direct contact with a Census collector at collection (Mailback) or indirect through a reminder card to return the Census form (Non-contact).

This report shows that posted-in forms can be linked to lower data quality when compared to collected forms. The only real impact on non-response rates is likely to be apparent for the Religion, Income, and Structure of dwelling variables although lower data quality may be a problem at the small area level in capital cities.

The greatest impact of the mailback option, however, is due to the low return rate of forms left for posting back. On dummy forms (representing forms which were not posted-in) almost all responses are missing whereas the majority of responses appear to be intact on posted-in forms. The number of dummy records in the data has more than doubled between the 1986 and 1991 Censuses. Further, mailback (MB) and non-contact (NC) dummy forms

represent over 90% of all dummies. In 1991 dummy records appear to have contributed about 0.9 percentage points more of the recorded non-response than in 1986 for variables requiring an answer from all persons. The impact on question non-response due to dummy forms varies according to the overall level of non-response and whether missing responses for specific variables are imputed.

From a field perspective, there are also some interesting points to note. The most important is that it appears that it is becoming slightly more difficult to collect completed forms from householders, necessitating the increased use of dummy forms. The total number of dummy forms raised, and the number remaining in the data after processing, has increased considerably since the 1986 Census. The return rate of posted-in forms is particularly low for forms resulting from non-contact situations. There may be some scope to improve the request to return card which is left when contact is not made at collection, or to introduce some type of reminder system. Finally, some collectors failed to mark the Structure of dwelling question before delivery, thereby contributing to the not stated rate. The non-response rate for final data for this variable indicates that some collectors may be failing to mark this question even after collection.

Any expanded promotion of the availability of the mailback option in future Censuses would inevitably lead to a greater use of posted-in forms. Even without this, and with the use of effective publicity campaigns aimed at increasing the level of cooperation in the Census in the community, the level of posted-in forms may continue to increase in future Censuses.

An important issue to consider is whether there would be an increase in the number of mailback dummies or non-contact dummies raised, or both, because these situations appear to have different consequences for data quality. The impact of an increase in mailback dummies on non-response rates would probably not be great (except for Income and Religion as discussed above). However any increase in the number of non-contact dummies is a concern - in the 1991 Census only about 33% of forms resulting from NC situations were posted-in.

In conclusion, there is a need to continue to monitor the impact of the mailback option on data quality, given the scope for these procedures to have a significant and detrimental impact on the data. Associated with this is a need for reliable and timely data on the rates of return, the incidence of posted-in forms, and not stated rates for posted-in forms.

## **APPENDIXES**

### **APPENDIX 1:**

#### **Problems associated with the large number of posted-in forms received as identified by DPC and Census Field Organisation staff**

- Space constraints and understaffing - estimates were for about 60,000 - 75,000 mailed forms based on the 1986 Census experience. The actual number received was about two to three times higher (see Section 3 for details).
- About one fifth of the envelopes received were non-standard, with indications from respondents suggesting Collectors had either run out or were reluctant to give out envelopes. As most of these envelopes were small the Census forms were usually folded and creased, often resulting in the need for transcription to permit the forms to be fed into the OMR machines.
- The large number of forms arriving too late to be inserted in Precapture (details in Section 3) made it necessary to insert the bulk of these forms during Main Processing,

a very disruptive exercise.

- The National Evaluation Conference on the Collection Phase of the Census expressed concern about the increase in mailback follow-up and the increase in loss of quality control by management during collection.
- It was also pointed out at this conference that maybe there was too much emphasis on refusals but too little on mailback forms which could be of such poor quality to be little better than disguised refusals.
- There were problems with the mailback labels, some forms were posted to the State Offices, others to the DPC. This created extra work in re-directing the mail.

## **APPENDIX 2:**

### **Back of Census form**

HARD COPY AVAILABLE

## **APPENDIX 3:**

### **'Request to Return' Card**

HARD COPY AVAILABLE

## **APPENDIX 4:**

### **Posted-in form counts (1986/1991 Censuses) - Sources**

In 1986, a tally form was used by Special Projects to count the mail being received each day. One of the categories of this form was 'form completed/blank' (completed meaning not totally blank or containing some written information). According to this tally form, a total of 54,571 posted-in forms (51,320 completed and 3,251 blank) were received. This figure does not include the number of envelopes received that were thought to be refusals which were later found to be non-refusal cases. Thus the final count of posted-in forms was revised to 57,261, an increase of 2,690. This revised figure needs to be kept in mind, especially for the comparison of inserted posted-in forms by State between 1986 and 1991 in Table 6. There are no indications how many of those additional forms were blank.

Checking the mailback codes on the Special Evaluation Flag of the 1986 FURF the total number of posted-in forms was found to be 48,748. Indications are that tally form counts were more reliable than FURF counts. Due to organisational problems towards the end of processing, not all forms had their Special Evaluation Flag recorded on the system. Also, the reliance on key-entry and the potential for clerical error must be considered in this context. For these reasons, 1986 data were based on the Special Projects tally forms.

In 1991, a tally form was again used at the DPC to record details of forms mailed in. The mail processing section used a category 'form complete/blank' similar to the 1986 Census. There appeared to be some confusion by clerical staff about the exact definition of 'blank'. Accordingly, there can only be estimates about the number of blank forms. Manual counts by mail processing staff found the number of posted-in forms received to be approximately 175,000. This figure included about 15,000 forms that were not inserted during the early stages of the process, either because they arrived too late to be inserted during Precapture (mainly forms from NSW and VIC, the States processed in the first month of DPC operation)

or because they were blank or contained obviously untrue or not identifiable information. Of these 15,000 forms, 11,000 were inserted during a later processing stage, Main Processing. This means a total of around 4,000 forms were not inserted in CD packs. Indications are that the majority of these forms were blank. Some forms from the smaller States that were processed from the second month of operation (like TAS or WA) were not inserted because they arrived too late, and some forms were not inserted because they contained obviously untrue or not identifiable information.

The number of inserted posted-in forms, derived from running a SAS program against the posted-in form indicators on the IFURF, was found to be 171,199. Adding the number of forms not inserted gives a total of 175,199 posted-in forms received. This figure relates closely to the 175,000 forms tallied by mail processing staff. 1991 data were based on this IFURF data.

#### **APPENDIX 5:**

#### **Occupied private dwellings Occupied private dwellings, 1986 and 1991 Censuses, by State**

State	Occupied private dwellings		Percentage increase
	1986 Census	1991 Census	%
New South Wales	1,832,642	1,987,265	8.4
Victoria	1,356,235	1,475,393	8.8
Queensland	860,813	1,017,802	18.2
South Australia	475,987	515,705	8.3
Western Australia	467,264	549,931	17.7
Tasmania	149,458	163,001	9.1
Northern Territory	42,556	50,542	18.8
Australian Capital Territory	79,561	92,880	16.7
<b>Australia</b>	<b>5,264,516</b>	<b>5,852,519</b>	<b>11.2</b>

This table supplements Table 7 where the proportion of posted forms as a percentage of occupied private dwellings is shown. Occupied private dwellings include caravans in caravan parks for both 1986 and 1991 Censuses.

#### **APPENDIX 6:**

#### **Comparison of not stated rates for posted-in and collected forms**

HARD COPY AVAILABLE

NOTE: Not stated rates were not available for the following variables:

- Marital status
- Qualification (highest) field)
- Labour force status
- Rent (weekly)

#### **APPENDIX 7:**

#### **Comparison of not stated rates for posted-in forms received early and posted-in forms received late**

HARD COPY AVAILABLE

#### **APPENDIX 8:**

#### **Comparison of not stated rates for mailback and non-contact posted-in forms**

HARD COPY AVAILABLE

Please note: Only variables with a tally of more than three not stated codes per variable (mailback or non-contact forms) were included due to the very low incidence of non-contact posted-in forms.

## **About this Release**

### **ABOUT THIS RELEASE**

In the first few weeks of processing staff at the DPC were confronted with the arrival of an unprecedented number of Census forms by mail. The sheer volume of forms arriving every day greatly exceeded all expectations and posed a variety of logistical problems for management. The main concern expressed was about the impact on overall data quality as mailed forms were felt to generally include poorer quality data. This experience also showed the need to analyse the issue of posted-in forms in detail to enable the development of strategies for future Censuses.

---

---

© Commonwealth of Australia

All data and other material produced by the Australian Bureau of Statistics (ABS) constitutes Commonwealth copyright administered by the ABS. The ABS reserves the right to set out the terms and conditions for the use of such material. Unless otherwise noted, all material on this website – except the ABS logo, the Commonwealth Coat of Arms, and any material protected by a trade mark – is licensed under a Creative Commons Attribution 2.5 Australia licence